



CIRRELT

Centre interuniversitaire de recherche
sur les réseaux d'entreprise, la logistique et le transport

Interuniversity Research Centre
on Enterprise Networks, Logistics and Transportation

Separating Moral Hazard from Adverse Selection and Learning in Automobile Insurance : Longitudinal Evidence from France

Georges Dionne
Pierre-Carl Michaud
Maki Dahchour

September 2010

CIRRELT-2010-40

Bureaux de Montréal :

Université de Montréal
C.P. 6128, succ. Centre-ville
Montréal (Québec)
Canada H3C 3J7
Téléphone : 514 343-7575
Télécopie : 514 343-7121

Bureaux de Québec :

Université Laval
2325, de la Terrasse, bureau 2642
Québec (Québec)
Canada G1V 0A6
Téléphone : 418 656-2073
Télécopie : 418 656-2624

www.cirrelt.ca

Separating Moral Hazard from Adverse Selection and Learning in Automobile Insurance: Longitudinal Evidence from France

Georges Dionne^{1*}, Pierre-Carl Michaud², Maki Dahchour³

- ¹ Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation (CIRRELT) and Canada Research Chair in Risk Management, HEC Montréal, 3000, Côte-Sainte-Catherine, Montréal, Canada H3T 2A7
- ² Department of Economics, Université du Québec à Montréal, C.P. 8888, succursale Centre-Ville, Montréal, Canada H3C 3P8
- ³ Aviva Canada National Personal Lines, 2206 Eglinton Avenue East, Scarborough, Ontario M1L 4S8

Abstract. The identification of information problems in different markets is a challenging issue in the economic literature. In this paper, we study the identification of moral hazard from adverse selection and learning within the context of a multi-period dynamic model. We extend the model of Abbring et al. (2003) to include learning and insurance coverage choice over time. We derive testable empirical implications for panel data. We then perform tests using longitudinal data from France during the period 1995-1997. We find evidence of moral hazard among a sub-group of policyholders with less driving experience (less than 15 years). Policyholders with less than 5 years of experience have a combination of learning and moral hazard, whereas no residual information problem is found for policyholders with more than 15 years of experience.

Keywords. Moral hazard, adverse selection, learning, dynamic insurance contracting, panel data, empirical test.

Acknowledgements. Earlier versions of the paper benefited from comments made by Jaap Abbring, Charles Bellemare, Denis Bolduc, Jean-Yves Duclos, Sylvain Lassare, Anne Lavigne, Jean Pinquet, Bernard Salanie, Arthur van Soest, three anonymous referees and numerous seminar participants. We also thank Béatrice Gatterer from the Fédération Française des Sociétés d'Assurances (FFSA) for providing us with data on insurance contracts in France. Financial support by INRETS, Social Sciences and Humanities Research Council of Canada (SSHRC) and FFSA is acknowledge.

Results and views expressed in this publication are the sole responsibility of the authors and do not necessarily reflect those of CIRRELT.

Les résultats et opinions contenus dans cette publication ne reflètent pas nécessairement la position du CIRRELT et n'engagent pas sa responsabilité.

* Corresponding author: Georges.Dionne@cirrelt.ca

Dépôt légal – Bibliothèque et Archives nationales du Québec,
Bibliothèque et Archives Canada, 2010

© Copyright Dionne, Michaud, Dahchour and CIRRELT, 2010

1. Introduction

Information asymmetries generally follow two distinct pathways. Adverse selection strongly predicts a positive correlation between the accident probability of a policyholder and the generosity of his insurance contract. In the presence of moral hazard, the positive correlation is caused by the unobservability of effort to prevent accidents. Generous coverage reduces the expected cost of an accident and therefore reduces the incentives for safety. In the end, both pathways predict a positive correlation between accidents and coverage within a risk class. This suggests an empirical test for asymmetric information often referred to as the conditional correlation test.

The evidence is not conclusive concerning the existence of residual asymmetric information in automobile insurance markets.² Some studies using the conditional correlation approach on cross-sectional data find evidence of asymmetric information (Dahlby, 1992; Puelz and Snow, 1994; Richaudeau, 1999; Cohen, 2005; Kim et al., 2009) while others did not (Chiappori and Salanie, 2000; Dionne et al., 2001). One major criticism of the conditional correlation approach with cross-sectional data is that it does not allow separation of adverse selection from moral hazard (Chiappori, 2000).³

Abbring et al. (2003) investigate the dynamics in claims as a way of directly testing for moral hazard. Under most experience rating regimes, an at-fault claim raises the marginal future cost of another claim. Hence, such regimes should promote safe driving, at

² Results in other markets can be found in Chiappori et al. (1998), Dionne and St-Michel (1991), Fortin and Lanoie (1992), Hendel and Lizzeri (1999, 2003), Finkelstein and Porteba (2004), Finkelstein and McGarry (2006). See Cohen and Siegelman (2009) for a recent survey.

³ Also, with cross-sectional data, unobserved confounders such as risk aversion may mask evidence of asymmetric information. Chiappori et al. (2006) propose cross-sectional tests based on profit maximization in competitive markets that are robust to heterogeneity in preferences.

least in theory. Empirically, negative temporal occurrence dependence in accidents within a risk class should be observed under moral hazard. The authors find little evidence of moral hazard in France.⁴ In fact, there is only weak evidence among inexperienced drivers, which points to learning rather than moral hazard, i.e. beginners who learn they are bad risks exert caution.⁵ To separate learning leading to adverse selection (asymmetric learning) from moral hazard, we consider the case where information on contracts and accidents is available for multiple years in the form of panel data. We exploit dynamics in accidents and insurance coverage controlling for dynamic selection due to unobserved heterogeneity. Changes in insurance coverage allow us to construct two additional tests, which, coupled with the negative occurrence test of Abbring et al. (2003), allow us to separate moral hazard from asymmetric learning.

We analyze the identification of asymmetric learning and moral hazard within the context of a tractable structural dynamic insurance model. From the solution of the model, we simulate a panel of drivers behaving under different information regimes or data generating processes (with or without both moral hazard and asymmetric learning). We validate our empirical tests on simulated data generated from these different information regimes. We then apply these tests to longitudinal data on accidents, contract choice and experience rating for the period 1995-1997 in France. We find no evidence of information

⁴ The French experience rating system rates at-fault claims (Dionne, 2001; Picard, 2000; Pinquet, 1999). The information is public and shared across insurers. The system is enforced by a law stipulating the penalty (malus) and reward (bonus) in case of an at-fault accident. The rating coefficient, or the bonus-malus coefficient, is applied proportionally to the base premium at the time the signing/renewal of the contract. On the effectiveness of experience rating, see Dionne and Doherty (1994) and de Garidel-Thoron (2005).

⁵ Their estimate of negative occurrence dependence among inexperienced drivers is insignificant. Further, combining claims at-fault and other claims changes the result significantly, as would be predicted under asymmetric learning (all accidents should matter). Dionne et al. (2005) and Israel (2004) apply similar tests in other countries (Canada and U.S. respectively) and find evidence of moral hazard. Abbring et al. (2008) analyze accidents and claims reporting in a dynamic setting with moral hazard using data from the Netherlands. They find evidence of moral hazard.

problems among experienced drivers (more than 15 years of experience). For drivers with less than 15 years of experience, we find strong evidence of moral hazard but little evidence of asymmetric learning. We obtain evidence of asymmetric learning, despite the small sample size, when focusing on drivers with less than 5 years of experience.

The remainder of the paper is structured as follows. In Section 2, we present the theoretical model we use to construct empirical tests. In Section 3, we present our empirical tests and validate these data using simulated data from the theoretical model in Section 2. In Section 4, we present results of the tests applied to French panel data. Section 5 concludes.

2. Theoretical Model

To investigate how contract choices and accident outcomes allow us to distinguish moral hazard and asymmetric learning leading to adverse selection, we start with a dynamic model similar to Abbring et al. (2003). They build a model where policyholders directly choose the probability of future accidents by exerting effort. The marginal benefit on the accident probability and expected future premium is weighted against the immediate utility cost of effort. Given the time frame of their data, they focus on the dynamics within the contract-period in claims. They show that under an experience rating system such as the one used in France, an accident not only raises the average cost of future accidents but also the marginal cost. Given that premiums and deductibles are low relative to income and wealth, such that wealth effects do not dominate, they find that the optimal level of effort increases with past accidents and the corresponding increase in the premium. Hence, at the time of an accident, incentives to exert effort jump discontinuously. They then derive an

empirical test based on this prediction.

In order to consider other information problems in a multi-period context, we extend their model along two dimensions. First, we add a contract choice decision to the model. We also consider the possibility that there is learning (drivers and insurers learn about an individual's innate risk). Over time, this may lead to adverse selection if the policyholder learns faster than the insurer, or to full information if both share symmetric learning. Because the data we will use contains all accidents (not only claims), we allow consumers to learn faster than insurers, hence allowing for asymmetric learning. A final distinction is that we use a discrete-time model because our data are annual. We discuss timing issues below. Adding these extensions makes analytical solutions difficult, so we solve the model numerically and show how the dynamics in accidents and insurance coverage allow us to separate moral hazard from adverse selection and learning. Because the model has clear policy parameters that govern the presence of moral hazard, adverse selection and learning, we can simulate a cohort of drivers from different scenarios and confirm whether the empirical tests we propose can separate these information problems. We now give a brief overview of the model and its key elements.

The most important insurance decision in France is that of buying comprehensive insurance coverage (CC) in addition to compulsory "*responsabilité civile*" or limited liability coverage (LL). By law, every driver must purchase an LL contract that protects third parties if the driver is at fault. However, the LL contract does not cover the driver's damages if he is at fault. The CC contract covers such damages and the policyholder pays a deductible only when he is at fault, which varies across contracts. Finally, the other party's insurer pays all damages if the policyholder is not at fault. The insurer observes only the

claims, whether the insured is at fault or not, while the insured observes all accidents.

The total premium paid is experience rated. It is scaled by a coefficient, the bonus-malus coefficient, which is a function of the history of claims where the driver is at fault. The individual bonus-malus coefficient is publicly known by all insurers and the premium adjustment is mandated by law, so insurers are fully committed to the bonus-malus pricing scheme. New drivers start with a bonus-malus coefficient of 1 and the coefficient varies between 0.5 and 3.5. Each at-fault accident increases the malus by 25%, while a year without accidents leads to a 5% reduction.

We assume drivers differ in terms of risk type (or ability). In the model, agents first obtain insurance without knowledge of their own risk. They learn about risk with experience and accidents. Accidents differ depending on whether the driver is at fault or not.⁶ Although the insurer observes the bonus-malus he does not learn as fast as the agent about his riskiness, partly because it observes only the claims. Thus asymmetric learning develops, which may lead to pure adverse selection in contract choices (Rothschild and Stiglitz, 1976; Wilson, 1977; Cocker and Snow, 1986). The agent can influence his accident probability by exerting effort. Effort is unobservable to the insurer and there is moral hazard within a given risk class because the driver has less incentive to exert effort under more generous insurance contracts (Holmstrom, 1979; Shavell, 1979; Winter, 2000) or when past accidents increase the marginal cost of future accidents (Abbring et al., 2003).

Contracts are renewed annually. An agent makes two decisions at each period $t = 1, \dots, T$ consisting of the choice of contract and of the level of effort to prevent

⁶ We do not model the decision to report an accident or not. We assume all at-fault accidents are reported which is reasonable because these accidents typically involve two parties.

accidents. For the sake of simplicity we assume time represents driving experience. The timing of the decisions is shown in Figure 1.

After the agent has made his decisions, uncertainty is resolved. We assume only one accident can occur in each period (few drivers have more than one accident per year). We denote by $n_t = \{0, 1, 2\}$ the occurrence of accidents where 0 means no accidents, 1 denotes an accident where the driver is not at fault and 2 an accident where he is at fault. As usual n_t is unknown prior to making decisions at t . At the beginning of the following period damages are paid prior to the agent's making any new decision. Finally, the agent pays the premium on the new contract when he makes his contract decision, conditional on past accidents and beliefs.

2.1. Insurance Contracts

We denote the choice of the CC contract as $d_t = 1$; $d_t = 0$ when only the LL contract is retained as coverage. An accident results in a fixed monetary loss, L . If $d_t = 1$ and if the policyholder has an at-fault accident, he pays, at most, the deductible $f < L$. If he does not purchase CC, he pays damage L . If the driver is not at fault, no payments are made by the driver, regardless of whether CC is purchased.

The premium has two components: an *a priori* and an *a posteriori* pricing component. The *a posteriori* component is a function of the driver's accident history, summarized by his bonus-malus b_{t-1} at the start of the preceding contract period, and the occurrence of an at-fault accident in the same period. The new bonus-malus for the current contractual year is given by

$$b_t = b(b_{t-1}, n_{t-1}) = \begin{cases} \delta_b b_{t-1} & \text{if } n_{t-1} \neq 2 \\ \delta_m b_{t-1} & \text{if } n_{t-1} = 2 \end{cases} \quad (1)$$

where δ_b is the bonus coefficient (0.95) and δ_m is the malus coefficient (1.25). In the French market, the first at-fault accident does not increase the premium if the policyholder has the minimum coefficient (0.5). We keep track of this particular clause in the model.

The *a priori* pricing component depends on the choice of the CC contract. The total premium paid (pr) is the product of the bonus-malus and the *a priori* component

$$pr_t(b_t, d_t) = b_t \exp(\rho_0 + \rho_1 d_t) \quad (2)$$

where $\exp(\rho_1) > 0$ denotes the percentage increase in the premium for the CC contract and $\exp(\rho_0)$ is the base premium for the LL coverage.

We assume it is costly to change coverage over time. The cost is given by ψ and is symmetric (the same for increasing and decreasing coverage). Contract choices are quite persistent in the data and not allowing for switching costs would entail too many transitions compared with the data. Furthermore, some studies show that price dispersion is large for the same insurance product across insurers, which is consistent with the existence of switching or search costs (see Schlesinger and von der Shulenburg, 1993).

2.2. Effort and Accidents

Agents can choose to exert prevention effort $e_t = \{0,1\}$ during the contractual year. Effort reduces the probability of both types of accident. Assume the probability of an accident takes the multinomial logit form

$$p(n_t = j | e_t, \alpha) = \frac{\exp(\mu_{j0} + \mu_{j1}e_t + \mu_{j2}\alpha)}{1 + \sum_{k=1,2} \exp(\mu_{k0} + \mu_{k1}e_t + \mu_{k2}\alpha)}, \quad j = 1, 2 \quad (3)$$

where μ_{j1} is the coefficient affecting the marginal productivity of effort for the probability of an accident of type j . The individual is not certain of his true innate risk classification α , which is time invariant. It can take two values $\alpha = \{0, \alpha_H\}$, where α_H denotes the high-risk type. These accident probabilities are serially uncorrelated over time, for a given α and effort sequence. We also assume for simplicity that there is no “experience” effect irrespective of the accident history, i.e. accident probabilities do not depend on time.

The experience of the policyholder can be used to construct expectations about his innate riskiness. He knows equation (3) and the fraction of high types α_H in the population, δ_H . We denote by $\pi_t(\alpha_H | n_1, \dots, n_{t-1}, e_1, \dots, e_{t-1})$ the subjective probability that the policyholder is a high-risk type for the current contractual period given his effort and accident realization in the previous periods. We denote the accident history by $n^t = (n_1, \dots, n_t)$ and do the same for other variables. Using Baye's rule, the probability that the driver is a high-risk type given his history or experience up to t is

$$\begin{aligned} \Pr(\alpha_H | n^{t-1}, e^{t-1}) &= \frac{\Pr(n^{t-1} | e^{t-1}, \alpha_H) \Pr(e^{t-1} | \alpha_H) \Pr(\alpha_H)}{\Pr(n^{t-1} | e^{t-1}) \Pr(e^{t-1})} \\ &= \frac{\Pr(n^{t-1} | e^{t-1}, \alpha_H) \Pr(\alpha_H)}{\Pr(n^{t-1} | e^{t-1})} \end{aligned} \quad (4)$$

given that effort is non-stochastic. Using equation (3), we can rewrite equation (4) as

$$\begin{aligned}
 \Pr(\alpha_H | n^{t-1}, e^{t-1}) &= \frac{\Pr(n_{t-1} | e_{t-1}, \alpha_H) \Pr(n^{t-2} | e^{t-2}, \alpha_H)}{\Pr(n_{t-1} | e_{t-1}) \Pr(n^{t-2} | e^{t-2})} \Pr(\alpha_H) \\
 &= \frac{\Pr(n_{t-1} | e_{t-1}, \alpha_H)}{\Pr(n_{t-1} | e_{t-1})} \Pr(\alpha_H | n^{t-2}, e^{t-2})
 \end{aligned} \tag{5}$$

Thus, denoting π_t as the subjective belief in period t , equation (5) simplifies to a recursive form

$$\pi_t = \frac{\pi_{t-1} \Pr(n_{t-1} | e_{t-1}, \alpha_H)}{\pi_{t-1} \Pr(n_{t-1} | e_{t-1}, \alpha_H) + (1 - \pi_{t-1}) \Pr(n_{t-1} | e_{t-1}, 0)}. \tag{6}$$

This equation shows how policyholders update their prior probability of being a particular risky type from π_{t-1} to π_t . Having an accident makes it more likely they are the high-risk type. It makes it even more likely if they exerted effort but still had an accident. Every period, the subjective probability of an accident of type j is given by

$$\tilde{p}(n_t = j | e_t, \pi_t) = \pi_t p(n_t = j | e_t, \alpha_H) + (1 - \pi_t) p(n_t = j | e_t, 0), \quad j = 1, 2 \tag{7}$$

2.3. Maximization Problem

Within a period, utility is assumed separable in consumption (c_t) and effort

$$u(c_t, e_t) = \frac{c_t^{1-\sigma}}{1-\sigma} + \gamma e_t \tag{8}$$

The coefficient of relative risk aversion is σ and the marginal disutility of effort is $\gamma < 0$.

Agents receive income y_t every period. We assume income is stochastic and follows a first order Markov process. Its distribution $F(y_t | y_{t-1})$ is known to the agent. We also assume there are no assets in the model so that

$$c_t = y_t - pr(d_t, b_t) - \psi I(d_t \neq d_{t-1}) - n_{2,t-1} L(d_{t-1}) \tag{9}$$

where $L(d_{t-1}) = (1 - d_{t-1})L + d_{t-1}f$ and $n_{2,t-1}$ is equal to 1 if an at-fault accident occurs and zero otherwise.⁷ The indicator function is given by $I(z) = 1$ if z is observed and zero if not. Finally, we denote the agent's subjective discount factor as β .

The deterministic part of an agent's state at the beginning of a period is given by $s_t = (\pi_{t-1}, b_{t-1}, d_{t-1}, e_{t-1})$. His/her optimization problem can be expressed as a series of one-period problems using Bellman's principle of optimality. We have

$$V_t(s_t, n_{t-1}, y_t) = \max_{e_t, d_t} u(c_t, e_t) + \beta \left[\int_{y_{t+1}} \sum_n V_{t+1}(s_{t+1}, n, y_{t+1}) p(n_t = n | e_t, \pi_t) dF(y_{t+1} | y_t) \right] \quad (12)$$

subject to the transition equations above for the bonus-malus, accidents and subjective probability (equations 1,2 and 6 respectively) and for $t=1, \dots, T$. The maximization problem can be solved by backward recursion from the last period $T=40$. We discretize the state-space for the log of the bonus-malus b , over the interval $[0.5, 2]$ and π over the unit interval.⁸ We use 35 grid points for both. We use bi-dimensional cubic spline approximation at each point in the state-space because the optimal solution and transition from one bonus-malus and subjective probability may fall between two points where we have calculated the value function in the following period. It is important to stress that the optimal solution for a given period will be a function of all variables at the beginning of

⁷ We do not allow for savings. This would complicate the numerical solution of the model substantially given that we have two continuous state variables. It is not clear whether wealth effects are important for this analysis. As Abbring et al. (2003) argue, premiums and deductibles are fairly low relative to household income and wealth in France. As we will see, even the mean loss for an accident on an LL contract is relatively low relative to income. See Chiappori et al. (1994) for a dynamic model of moral hazard with financial assets.

⁸ Given the parameters of the accident process, the maximum bonus-malus coefficient approaches 2. Very few observations in the SOFRES panel have a bonus-malus greater than 2 (4 observations out of 12,000).

the period. Hence, we will focus on the dynamics in contract choice and accidents to derive empirical tests of moral hazard and asymmetric learning.

2.4. Calibration

We first present the data. We then justify the calibration for the analysis. We calibrate the model in an attempt to get realistic profiles for the key observed variables. We do not claim to be able to identify all parameters from the data. Nevertheless, we attempt to “choose” reasonable parameters, either borrowed from the literature or calibrated using either SOFRES data or market data provided by the French Federation of Insurers (FFSA).

2.4.1. SOFRES Longitudinal Survey

The data are from the SOFRES longitudinal survey, *Parc Automobile*, which is a rotating panel representative of the French policyholders and their vehicles for the years 1995, 1996 and 1997. Respondents were interviewed by mail (questionnaires were sent each January) about several topics including car insurance and their accident history. The head of the household completed the questionnaire and returned it to SOFRES. SOFRES is an independent survey organization conducting monthly consumer surveys and is not an insurer. This allowed information gathering on both claims and accidents, including those not reported to insurers. Furthermore, SOFRES re-interviews individuals even if they switch insurers. Both of these characteristics of the survey are advantages from the point of view of testing for asymmetric information in insurance markets using longitudinal data. Existing studies often use data from one insurer (or a group of insurers). In that case, panel attrition is a potential problem because only “good customers” may remain in the panel. Accidents rather than claims are useful as accidents convey additional “private”

information about respondents that is not in the public domain. We rely on self-reports from respondents rather than actual administrative records of claims, accidents, and perhaps more importantly, contract characteristics. Finally, the timing of the survey does not correspond exactly to the renewal of insurance contracts. Contracts are renewed year round in France while the survey is conducted in January. Accidents over the previous 12 months are reported.

We define an observation as a respondent-vehicle pair. To match vehicles across waves, we use a combination of the respondent ID, the first 4 digits of the license plate and the year when the car was manufactured. A sizeable proportion of contracts are observed for less than three years. SOFRES aims to maintain the representativity of its sample over time, which means that some respondents are not re-interviewed for exogenous reasons or reasons related to observable characteristics that we control for (i.e. region). We keep entries from 1996 and 1997 even if they were not present in 1995. The appendix provides more details on the number of observations and the participation patterns in the survey.

The survey contains four modules. The first covers the socio-economic characteristics of policyholders. The second covers characteristics of the vehicle. In the appendix we provide descriptive statistics on these data. The third concerns insurance contracts. It provides the current bonus-malus coefficient and the current type of insurance coverage: CC or LL contract. The bonus-malus coefficient is updated at the end of every contractual year and is constant within a year (i.e. premiums can be adjusted only at the end of each contract year). The new bonus-malus appears only when it is time to negotiate a new contract. Insurers are committed to apply the regulated bonus-malus but are free to renegotiate all other parameters of the contract. Given the design of survey, this implies

that the bonus-malus in a given year does not take account of accidents that have taken place during the months preceding the survey date but after the contract renewal date. When regressing the bonus-malus coefficient on accidents, this could create obvious simultaneity, so we use the bonus-malus coefficient from the last year to avoid simultaneity. Finally, the survey covers accidents and claims, and as mentioned earlier, both are reported. However, we cannot distinguish at-fault claims from other claims.

In the top panel of Figure 2, we report the prevalence of CC contracts, accidents and claims, and the bonus-malus coefficient by experience (in years) in the SOFRES panel. Comprehensive coverage increases with experience, from roughly 40% for drivers with little or no experience to over 70% for those with more than 30 years of experience. This is despite the fact that less experienced drivers have more accidents and make more claims, which should lead them to choose more coverage. About one fourth of drivers with little or no experience have an accident (very few have more than one). This fraction decreases with experience. Roughly 68% of accidents are reported to the insurer. The under-reporting of accidents appears to be relatively unrelated to years of driving experience. In the bottom panel, we report the median, the 75th and 95th percentiles of the bonus-malus distribution by experience. Among beginners, the median bonus-malus is 0.75. It is well documented that many new drivers start with a favorable bonus, potentially inherited from their parents. Very few respondents have a bonus-malus greater than one (less than 3%). The vast majority of drivers with more than 20 years of experience have the maximum bonus of 0.5.

2.4.2. Calibration

We chose a baseline calibration such that both moral hazard and asymmetric learning are present. This is our baseline scenario. As mentioned above, we do not claim to be able to estimate all parameters, such as risk aversion, from the data. That full exercise, although interesting, is fairly ambitious and has been left for further research. We use parameter values that provide reasonable profiles comparable to those in the data. We have omitted the introduction of non-stationarity, which would be required to fit the model to the data. For example, premiums are likely to be a function of years of experience, as is the probability of having an accident. This could explain the declining accident profile observed in the data. Nevertheless our calibration exercise is likely sufficient for our purposes. Table 1 shows the parameters we use in the baseline scenario. All monetary amounts are in 2009 euros (thousands).

Preference Parameters: We use a coefficient of relative risk aversion equal to 1.5. We calibrate this parameter so that it matches the fraction of policyholders with comprehensive coverage in the SOFRES panel. The value of 1.5 is at the low-end of the values reported in Attanasio and Weber (1995) [1.49-3.39] but larger than the values found in Hurd (1989) [0.7,1.12]. Because the average loss and premiums are small relative to income in our application, a value of 1.5 is likely reasonable for our purposes. The marginal disutility cost of effort is difficult to calibrate. We have assumed a value of -0.0005. The discount factor was assumed to be 0.985. The main results of the simulation exercise are not affected by these three values.

Accident Process: We calibrate the intercepts of the accident process such that it matches the fraction of accidents in the data. However, we do not observe at-fault claims in

the SOFRES panel, only total claims, which include accidents where another party is at fault. Abbring et al. (2003) report that the annual rate of at-fault claims is 6.4% in their sample, which is also from France. We use this figure to calibrate the intercepts. We assume the effort coefficient is -0.5 for both types of accident probabilities (at-fault or not). This yields a sizeable effect on the probability of having an accident. Finally we assume arbitrarily that the factor coefficient of the high risk-type μ_{12} is 0.75 in the accident probability equation. The parameter μ_{22} is set to 1.6 in the at-fault accident probability equation. Hence, we assume that high-risk types are more likely to have at-fault accidents than low-risk types. We assume, quite arbitrarily, that 30% of the population is high-risk.

Market Data: The SOFRES data set contains little information on premiums and deductibles, so we use aggregate information obtained from the FFSA for the year 1997.⁹ The average deductible (f) in France was quite low, 192 euros. The average loss (L) was 2,439 euros. The average premium for the LL coverage was 184 euros. The coefficient ρ_1 is set at 0.7 for the CC contract. The CC premium is 370 euros, which is double that of the LL premium. The bonus factor is 0.95 and the malus factor 1.25. We assume the switching cost is 50 euros, which given a median annual household income (adjusted using an equivalence scale) of roughly 14,000 euros, represents roughly one full day's pay (assuming 2000 hours of work annually).

Income Process: The SOFRES panel contains information on income but the information is categorical, and the bins are not necessarily natural ones. Instead, we use data from the French subsample of the European Community Household Panel (ECHP) for the years 1995-1997. We use household net income along with OECD equivalence scales

⁹ Information for 1995, 1996 was similar.

to adjust for household composition.¹⁰ We discretize income into 10 categories (deciles) and use the midpoint within each category as the approximate value in the simulations. We then compute the transition matrix across these deciles to obtain an estimate of $F(y_{t+1} | y_t)$. We use ECHP-provided weights when computing this matrix. The resulting median net income is 14,319 euros and there is a fair amount of persistence in income over time.

2.4.3. Simulation of the Baseline Scenario

We solve the model for the optimal decision rules given this choice of parameters. We use $T=40$ as the maximum experience level. We then simulate 2000 drivers until they reach 25 years of experience. When simulating individuals, we randomly draw the initial bonus-malus, subjective probability of being high risk, risk type and income. We draw the bonus-malus from the SOFRES empirical distribution for drivers with less than 2 years of experience. The average bonus-malus is roughly 0.75. The initial subjective probability is drawn from $\Phi(a + c)$ where c is the normal standard and a is such that the average subjective probability is 0.3. This is a natural average prior given that 30% of the population is high risk. The risk type is drawn independently of all other variables. Since we know each driver's type, either low- or high-risk, we compute statistics by risk type and experience.

We report the results of the simulations in Figure 3. As one would expect, there is a clear distinction between the behavior of high and low risk drivers. Differences grow larger

¹⁰ The OECD equivalence scale used in the ECHP is the sum of weighted household members where the weight is one for the first adult member, 0.5 for subsequent adults (over 18 years old) and 0.3 for each child (under 18 years old).

with experience as both types learn what type of driver they are. This can be seen from the first panel reporting the subjective probability of being high type. As the high risk individuals have more accidents, they slowly become more convinced that they are high risk. The opposite occurs for low-risk types. Since the low risk types have less accidents, particularly at-fault accidents, their bonus-malus coefficient falls more rapidly towards 0.5. The bonus-malus coefficient of the high risk types remains high. In our model, the bonus-malus coefficient is an excellent indicator of risk. Because we did not introduce any non-stationarity in the accident process, the fraction with accidents and at-fault accidents is fairly stable over time. The gap between the two groups is important. Hence, an econometrician analyzing those data would find large unobserved heterogeneity in accidents.

The portion with CC coverage is initially very similar between the two groups. This is because both risk types do not know yet what kind of driver they are likely to be. As they become more convinced, the high-risk types purchase the CC coverage more rapidly than those who learn they are likely low-risk. Over time, a positive correlation emerges between risk and coverage, which leads to adverse selection, as predicted by asymmetric learning. The rank correlation between accidents and CC coverage increases from zero to 0.05 by the time drivers reach 15 years of experience. A substantial proportion of the population exerts effort, and high risk drivers eventually make a greater effort to reduce their probability of having an accident. In terms of the joint distribution of effort and coverage, we find that very few drivers choose not to make an effort and forgo CC coverage (less than 5%). Most drivers choose either one or both options.

3. Empirical Tests

To derive empirical tests that allow us to separate moral hazard from asymmetric learning and adverse selection, we first look at the form of the optimal decision rules and how they depend on state-variables. We then focus on how to apply these tests to the SOFRES Panel.

3.1. Optimal Decisions

Denote by $V_t^{jk}(s_t, n_{t-1}, y_t)$ the option specific utilities at time t for each combination of contract j and effort k given the state s_t, n_{t-1}, y_t . To derive tests for moral hazard and asymmetric learning, we will focus on the dependence of decision rules on the state-space variables.

The optimal effort level is a discrete function $e_t^*(s_t, n_{t-1})$. The decision rule can be written as

$$e_t^*(s_t, n_{t-1}) = I(\max(V_t^{11}, V_t^{01}) - \max(V_t^{10}, V_t^{00}) > 0)$$

It depends on the state $s_t = (\pi_{t-1}, b_{t-1}, d_{t-1}, e_{t-1})$ and the realization of an accident in the previous period. This function is not easily tractable analytically. Hence, we resort to numerical solutions.

Moral hazard is present if effort depends on contract parameters. The two contract parameters that vary between individuals are coverage choice and the bonus-malus coefficient that affects the premium. There is moral hazard if optimal effort depends on these characteristics.

Abbring et al. (2003) show within the context of their model that optimal effort

increases with the premium via the bonus-malus coefficient. As the coefficient increases, due to an accident where the driver is at fault, the marginal cost of a future accident increases. Hence, effort is greater when the premium is higher, or alternatively when the bonus-malus coefficient is higher. This can be seen in the first graph in Figure 4. We plot the net utility of exerting effort, $\max(V_t^{11}, V_t^{01}) - \max(V_t^{10}, V_t^{00})$ against the bonus-malus at the beginning of the period. The dotted line denotes the case where the individual just had an at-fault accident (change in the bonus-malus factor) while the solid line denotes the case where no accident occurred. It is clear that incentives to exert effort are higher when an at-fault accident occurs. The incentive generally increases as the bonus-malus coefficient increases up to a b_{t-1} of roughly 1.4.

The optimal effort level may also depend on the last contract choice d_{t-1} . This first occurs because the presence of switching costs implies that drivers who already have CC coverage are more likely to prefer staying with this coverage than to reduce their coverage and exert more effort on the margin. This type of state-dependence implies that the previous contract choice should be predictive of current accidents. A complementary mechanism to the bonus-malus is at work. Those who had an at-fault accident at $t-1$ incurred different losses depending on whether they were covered or not. This affects disposable income and the marginal utility of consumption. Thus, the trade-off between coverage and effort. Because an uncovered loss will lead to a larger marginal utility of consumption than a covered one, the uncovered driver will exert more effort at t . These two effects together create a situation where current accident probabilities depend on past contract choice.

This can be seen in the second graph in Figure 4. The solid line represents a case

where the driver had only LL coverage at $t-1$ while the dotted line represents the case where he/she also had CC coverage. In both these circumstances we consider the case where the individual had an at-fault claim ($n_{t-1} = 2$). For most of the bonus-malus range that matters in the data ($\text{bonus} < 1.25$), effort is less likely when CC coverage was purchased at $t-1$. This occurs because the marginal utility of consumption is higher when only the LL contract was purchased, because the loss was absorbed entirely by the policyholder. This form of moral hazard is in the spirit of Lambert (1983).

Finally, optimal effort may depend on past accidents through the updating of the subjective belief from π_{t-1} to π_t following the occurrence of n_{t-1} . If the occurrence of n_{t-1} leads to an upward shift in π_t , a driver may exert more effort to lower his probability of having an accident at t . The driver is essentially learning about his level of risk and modifies his behavior accordingly. Thus, learning induces a negative relationship between current accidents and past ones.

The optimal contract choice is a discrete function of the same variables. With asymmetric learning, those who have more accidents and learn that they are more likely to be high risks may choose to purchase more coverage for a given premium. Note however that if the driver is at fault, the premium will increase due to the bonus-malus. Thus, the effect is ambiguous in the case of at-fault accidents. For a given bonus-malus, the prediction is clear for other accidents. This can be contrasted with the case where there is symmetric learning. Assume that no-fault accidents are not informative for the risk type and that at fault accidents are equally observable by both parties. In that case, the insurer will propose an insurance pricing scheme (with and without commitment) that will generate full information coverage in each period at different prices for the different risk

types (Boyer, Dionne and Kihlstrom, 1989; de Garidel-Thoron, 2005). Therefore, a new at-fault accident under symmetric learning should not trigger an increase in insurance coverage. However, it is unlikely that full symmetric learning can be obtained under moral hazard because the insured observes his effort and the link between at-fault claims and effort. In the third graph of Figure 3, we plot the net utility of buying CC coverage as a function of the subjective prior π_{t-1} . We distinguish between two cases: one where the individual has an accident $n_t = 1$ and one where he does not $n_t = 0$. It is clear from Figure 3 that the net utility of purchasing CC coverage is greater when an accident occurs than when it does not.

3.2. Empirical Predictions

The SOFRES panel contains the following data on individuals $i=1, \dots, N$ and $t=1, 2, 3$.

$$\left\{ \left\{ n_{it}, b_{it}, d_{it}, x_{it} \right\}_{t=1,2,3} \right\}_{i=1, \dots, N}$$

where n_{it} is equal to 1 if any accident occurs, and zero otherwise (it includes both at-fault and other accidents, either declared or not). The SOFRES panel does not contain information that would allow us to identify at-fault claims; it only tells us whether or not there was a claim for an accident. Both b_{it}, d_{it} have the same definition as in the model and x_{it} is a vector of policyholder-vehicle characteristics. The bonus-malus coefficient at t does not include accidents since the last renewal of the contract.

Substituting the solution for optimal effort in the accident process gives us the following equation for accident probability

$$\Pr(n_{it} | n_{it-1}, d_{it-1}, b_{it}, \pi_{it-1}, e_{it-1}, \alpha_i, x_{it}) \quad (10)$$

Let us first assume all right-hand side (conditioning) variables are observable. Two tests of moral hazard are possible based on the discussion in 3.1. The first test involves the bonus-malus coefficient b_t . This is analogous to the Abbring et al. (2003) test in a discrete time framework. If there is moral hazard, then the probability of an accident should be lower if the bonus-malus coefficient is higher, or if an at-fault claim occurred last period. Consequently, our first test is the following:

Test MH1: *There is no evidence of moral hazard if b_t has no effect on the accident distribution. A negative effect is consistent with the presence of moral hazard.*

Because π_{it-1} is not observed by the insurer, we cannot confirm the presence of moral hazard by finding a negative effect with the MH1 Test. Theory would predict that π_{it-1} would have a negative effect on n_{it} conditional on α_i . Furthermore, since π_{it-1} is positively correlated with b_{it} , this would imply that the negative effect of b_{it} on n_{it} could be explained by learning rather than moral hazard. Finding evidence of a negative effect on MH1 is therefore consistent with moral hazard but does not establish its presence. This is essentially the conclusion of Abbring et al. (2003). They find a statistically insignificant negative occurrence dependence effect among inexperienced drivers that they tentatively attribute to learning. However, the absence of moral hazard and learning is found for more experienced drivers in their study. The fact that α_i is unobserved is problematic because by construction, it is positively correlated with the bonus-malus and with accidents. However, we can account for α as a random effect component provided we deal with the

left-censoring problem observed in the data. The problem arises in other tests as well and we will discuss it further in Section 3.3.

The second test involves looking at the relationship between d_{it-1} and n_{it} . Because of switching costs and the fact that optimal effort is higher when the driver was not covered and had an accident, moral hazard would predict a positive effect of past CC coverage on the occurrence of accidents. This leads to a second test for moral hazard:

Test MH2: *There is no evidence of moral hazard if d_{it-1} has no effect on accident probabilities. A positive effect is interpreted as evidence of moral hazard.*

The unobservability of π_{it-1} bias the effect of d_{it-1} on n_{it} towards zero. This is because $\text{cov}(d_{it-1}, \pi_{it-1}) > 0$ but π_{it-1} has a negative effect on n_{it} conditional on α_i . Accordingly, a positive effect under MH2 implies that moral hazard is present. A zero (or negative effect) does not allow us to confirm the presence of moral hazard.

As mentioned in the previous section, we can test for asymmetric learning done by using the optimal contract choice equation. We can estimate the following contract choice probability from the data

$$\Pr(d_{it} \mid n_{it-1}, d_{it-1}, b_{it}, \pi_{it-1}, e_{it-1}, x_{it}) \quad (11)$$

Given that we control for b_{it} , the remaining effect of n_{it-1} on d_{it} should be positive under asymmetric learning. At-fault claims may also have an effect on contract choice if the premium increases. This will generally not be the case when the driver is not at fault. The fact that π_{it-1} cannot be observed will only bias the coefficient on n_{it-1} upward. Symmetric learning leads to a non-positive effect of past accidents on contract choice.

Thus, the following test:

Test AL: *There is no evidence of asymmetric learning if n_{it-1} has no effect on d_{it} . A positive effect can be interpreted as evidence of asymmetric learning.*

The three tests together provide a strong framework for separating moral hazard from asymmetric learning. Table 2 summarizes the tests and the conclusions. Only one ambiguous case emerges. It involves finding a negative effect on MH1, no effect on MH2 and a positive effect on AL. From this, we can conclude that there is AL but not that there is MH. This is because rejection of MH1 is consistent with AL due to the fact that π_{it-1} is unobserved by the insurer. Because MH2 is biased towards zero due to AL, we cannot detect the presence of moral hazard.

3.3. Econometric Model

We built an econometric model that allows us to conduct all three tests jointly. We consider a joint model of the probability functions in equations (10) and (11). We rely on parametric models. We specify each equation as a dynamic binary choice model with pre-determined regressors and an error component structure. We let error terms be correlated between equation (10) and (11), as unobservability of π_{t-1} and α_i would suggest. The error component structure is appealing given the likelihood of serial correlation in contract and accident outcomes. For example, equation (10) shows that we condition on α , which is unobserved. This type of bivariate dynamic choice model has been applied to stock and mutual fund ownership (Alessie et al., 2004) and female labor force participation (Michaud and Tatsiramos, 2009). A key question is how to deal with the initial condition problem. Below we use the solution proposed by Wooldridge (2005). We will test the adequacy of the model on simulated data.

More specifically, we specify equation (10) as

$$\begin{aligned} n_{it} &= I(x_{it}'\beta_n + \phi_{nd}d_{it-1} + \phi_{nm}n_{it-1} + \phi_{nb}b_{it} + \varepsilon_{n,it} > 0) \\ i &= 1, \dots, N, t = 1, \dots, T \end{aligned} \quad (12)$$

where $\varepsilon_{n,it}$ has an error component structure $\varepsilon_{n,it} = \alpha_{ni} + v_{n,it}$. We specify a similar equation for contract choice

$$\begin{aligned} d_{it} &= I(x_{it}'\beta_d + \phi_{dd}d_{it-1} + \phi_{dn}n_{it-1} + \phi_{db}b_{it} + \varepsilon_{d,it} > 0) \\ i &= 1, \dots, N, t = 1, \dots, T \end{aligned} \quad (13)$$

where again $\varepsilon_{d,it} = \alpha_{di} + v_{d,it}$. In both (12) and (13), all right-hand side variables are assumed to be predetermined such that they are independent of $v_{j,it}, j = d, n$. The bonus-malus variable b_{it} is a deterministic function of past accidents $b(n_{i0}, \dots, n_{it})$. We allow the pair of unobserved heterogeneity terms $\alpha_i = (\alpha_{ni}, \alpha_{di})$ to be jointly normally distributed with correlation ρ_α and variances $\sigma_{\alpha_j}^2, j = n, d$. We do the same for the residual error terms $v_{it} = (v_{n,it}, v_{d,it})$. Conditional on the unobserved heterogeneity terms, omitting the conditioning on x_{it} , and writing $(d_{it}, n_{it}) = z_{it}$, we have

$$\Pr(z_{i1}, \dots, z_{iT} \mid \alpha_i, z_{i0}, b_{i1}) = \prod_{t=1}^T \Pr(z_{it} \mid z_{it-1}, b_t, \alpha_i) \quad (14)$$

which is the probability of the joint sequence of accident and contract choices given the initial bonus-malus, contract choice and accident outcome. The period $t = 1$ will generally not be the start of the process. For example, the SOFRES panel does not follow individuals from the first year they have an insurance contract. Accordingly, we face a problem of left censoring. Equation (14) makes the problem clear. Integrating (14) over α_i implies that we

either need to assume that α_i is orthogonal to z_{i0}, b_{i1} , that z_{i0}, b_{i1} have a degenerate distribution, or that we know the joint distribution of $(\alpha_i, z_{i0}, b_{i1})$. It is implausible to assume orthogonality and to assume a degenerate distribution. Even for beginners, the bonus-malus appears to vary substantially. Many beginners start with a bonus of 0.5. As a result, we need to specify features of the joint distribution of $(\alpha_i, z_{i0}, b_{i1})$. Two solutions have been proposed, one by Heckman (1981) that consists of specifying $\Pr(z_{i0}, b_{i1} | \alpha_i)$ and integrating over the marginal of α_i . The other solution proposed by Wooldridge (2005) is to specify instead $\Pr(\alpha_i | z_{i0}, b_{i1})$ and maximize likelihood using probabilities in (14) conditional on z_{i0}, b_{i1} . In particular, Wooldridge assumes that $E(\alpha_{ji} | z_{i0}, b_{i1}) = \pi_j z_{i0} + \delta_j b_{i1}$, $j = d, n$. Defining $\eta_{ji} = \alpha_{ji} - E(\alpha_{ji} | z_{i0}, b_{i1})$ as the residual term, one can substitute $z_{i0}\pi_j + b_{i1}\delta_j$ in the index of (12) and (13) and then integrate (14) over the distribution of $\eta_i = (\eta_{di}, \eta_{ni})$. We assume the distribution is bivariate normal. The maximum likelihood estimator for all parameters of the model becomes

$$\theta_{ML} = \arg \max_{\theta} \frac{1}{N} \sum_{i=1}^N \log \int \Pr(z_{i1}, \dots, z_{iT} | \eta_i, b_{i1}, z_{i0}, x_{i1}, \dots, x_{iT}) dF(\eta) \quad (15)$$

We compute the two-dimensional integral in (15) by simulation, replacing the integral with the simulator

$$\frac{1}{R} \sum_{r=1}^R \Pr(z_{i1}, \dots, z_{iT} | \tilde{\eta}_{i,r}, b_{i1}, z_{i0}, x_{i1}, \dots, x_{iT}) \quad (16)$$

where $\{\tilde{\eta}_{ir}\}_{r=1}^R$ are draws from the bivariate normal distribution with parameters $(\sigma_{\eta m}^2, \sigma_{\eta d}^2, \rho_{\eta})$. The resulting estimator is the maximum simulated likelihood estimator. It is consistent for N, R going to infinity and asymptotically equivalent to ML for R / \sqrt{N} tending to zero. We use 50 draws along each dimension. We use the BFGS numerical maximization algorithm and compute robust standard errors with the sandwich estimator.

The MH1 test translates into a test of whether $\phi_{nb} < 0$. A higher bonus-malus, conditional on dynamic selection due to unobserved heterogeneity, gives an incentive to exert more effort and thus reduce accident probabilities. The MH2 test translates into a test of whether coverage in the last period increases accident probabilities in this period; it is a test of whether $\phi_{nd} > 0$ or not. Finally, a test of asymmetric learning is a test of whether an accident last period, conditional on the bonus-malus, leads to an increase in coverage this period. It is a test of whether $\phi_{dn} > 0$ or not.

3.4. Tests on Simulated Data

The fact that the solution for effort and contract choice is assumed to be linear in the index of (13) and (14) may lead to misspecification. We verify the performance of each test (MH1, MH2 and AL) on simulated data using the model from Section 2. This is particularly useful because some variables, like subjective beliefs and risk type, are unobserved in the data but observed in the simulated data set. We simulate a cohort of drivers and then keep observations from experience years 6 to 15 (10 years). We do not take observations from the start of the process precisely so that we have left-censoring, i.e.

initial outcomes at year 5 are correlated with risk type. This replicates what we have seen in the data. We stop at year 15 because few contract changes occur after this date and we already have 10 observations per driver in the estimate. We include income and experience (linear) as controls in the vector x_{it} .

We consider four scenarios. First, we consider the case where both moral hazard and asymmetric learning are present. We estimate the econometric model presented in section 3.3 using the simulated drivers. We then turn off each information problem sequentially to eventually reach a scenario where there is no asymmetric information. To turn off moral hazard, we set the productivity of effort in the accident and at-fault accident process equal to zero. This produces a solution where effort is zero. To reach a situation within the context of the model where we turn off asymmetric learning, we set the factor loading on the high-risk type in the accident process to zero. Hence, accidents where the driver is not at fault, which are not observed by the insurer in the model, do not convey information about the driver's risk type. Consequently, the bonus-malus contains all the information regarding the risk type if there is no moral hazard. Two caveats are in order for this simulation. First, the bonus-malus coefficient is not necessarily as optimal as the actuarial forecast based on the claims history. Also, the driver remains slightly more informed since he observes his own effort. Note that when we consider the scenario where both information problems are disregarded, the insurer learns as fast as the driver. Table 3 reports the results of the three tests for all four scenarios.

When we assume both moral hazard and asymmetric learning, all three tests capture the information problems. This is shown in column 1 of Table 3. Lag accidents have a positive and statistically significant effect on the procurement of CC coverage as predicted

by asymmetric learning conditional on the bonus-malus. Both moral hazard tests, MH1 and MH2, reveal the presence of moral hazard. The current bonus-malus has a negative effect on accidents while the lag of the contract choice has a positive effect. Column 2 reports what happens when we assume symmetric learning rather than asymmetric learning. In that case, the AL test does not pick up a residual effect of lag accidents on contract choice. This implies that within the context of the model the bonus-malus is as good as the subjective belief of the driver and corrects the premium such that no increase in coverage occurs. Both moral hazard tests still detect moral hazard. The third column shows what happens when we simulate data with asymmetric learning but no moral hazard. In that case the AL test picks up asymmetric learning while both MH tests do not identify moral hazard. The MH2 test even shows that lag contract choice is negatively correlated with accidents. The last column shows the results when we assume that there is no information problem; in that case, all tests yield negative conclusions. The AL test yields a negative effect of lagged accidents on current contract choice. Overall, the tests on simulated data appear to pick up the information problems assumed in the model despite the restrictive functional form assumed and the unobservability of some of the state variables.

Next, we verified whether our solution to control for unobserved heterogeneity and initial conditions is adequate. Although results in Table 3 are generally positive, we demonstrate what happens if we control directly for the risk type (because in the simulated data we know who is high risk) and if we do not control for the risk type. We do the same adding the lag subjective probability π_{it-1} . Finally, we consider the case where we do not control for unobserved heterogeneity. We report the results of the tests in Table 4 using data generated with moral hazard and asymmetric learning.

Results in the first column of Table 4 replicate those of column 1 of Table 3 using the Wooldridge initial condition solution with random effects. The second column reports results where we control for the risk type α_i directly. In this case, we do not need to control for the initial conditions z_{i0}, b_{i1} either. The results are quite close to those in column 1, suggesting that our method for controlling for initial condition and unobserved heterogeneity is adequate. Because π_{it-1} is also unobserved, we produce results in column 3 where we include it as a control. Again the conclusions do not change. The last column shows what happens if we do not account for initial conditions and unobserved heterogeneity. For the AL and MH2 tests, results are relatively robust. However, the MH1 test reveals an insignificant effect of the bonus-malus on accidents despite the presence of moral hazard in the data. This is because the bonus-malus is significantly positively related to risk type. If the bonus-malus is omitted, we have an upward bias in the coefficient as suggested by Abbring et al. (2003).

4. Results on SOFRES Panel

We then apply our tests to the SOFRES panel. We estimate separate models for two groups: those with less than 15 years and more than 15 years of experience in the first survey year considered.¹¹ We consider the first sample (less than 15 years) the “inexperienced” sample and the second (more than 15 years) the experienced sample. One might suspect that asymmetric learning is more likely among drivers with less

¹¹ Abbring et al. (2003) split their sample based on year of birth (before or after 1980). Because their analysis takes place during the period 1987-1989, this means that their inexperienced sample is composed of individuals with less than seven years of experience. Due to the relatively small sample size in our analysis, we use a larger experience group (more or less than 15 years).

experience.¹² This may be the result of the bonus-malus scheme because in the long-term the coefficient will capture unobserved risk types through at-fault accidents. Table A.3 in the Appendix gives descriptive statistics on transitions in the panel stratified by experience.

In Table 5, we report a summary of the estimate results for the two groups.¹³ We also consider a specification where we allow for a separate MH2 and AL effect for drivers with less than 5 years of experience (the beginners).

We do not find evidence of learning among the inexperienced group overall. The point estimate on lag accidents in the contract equation is positive but statistically insignificant (column 1). However, when we allow for a different asymmetric learning effect for the beginners, we find a large positive effect (0.863) that is barely statistically significant at the 5% level (p-value = 0.053, column 3). We do not find evidence of asymmetric learning among more experienced drivers (column 2). The point estimate is negative (-0.223) and statistically insignificant (p-value=0.483). This suggests that asymmetric learning vanishes relatively quickly.

The MH1 and MH2 tests generally agree for both groups. We find evidence of moral hazard among inexperienced drivers (less than 15 years, column 1). The point estimate is strongly negative on the bonus-malus (-2.241, p-value=0.081). This suggests that accidents at fault at $t-1$ trigger incentives to exert caution, resulting in fewer accidents at t . The MH2 test tells a similar story. The point estimate is positive (0.651) and statistically significant (p-value=0.026), as the theory would predict. We do not find stronger evidence of moral hazard among beginners (column 2). The effect for those with

¹² We do not exploit differences between claims and unreported accidents in this analysis. Although this decision could be endogenous, we do not have enough data on deductibles and losses to exploit this distinction effectively.

¹³ Complete results are available upon request.

less than 5 years of experience is no different from that of the rest of the group (point estimate is 0.17, p -value=0.351). There is no evidence of moral hazard among the more experienced group. The point estimate is 0.096 on the MH1 test with a p -value of 0.892. The same conclusion is reached using the MH2 test.

5. Discussion and Conclusion

In this paper, we have analyzed the identification of moral hazard from asymmetric learning within the context of a structural dynamic insurance model. We extended the model developed by Abbring et al. (2003) to include learning and insurance coverage choice. We derived two tests in addition to the negative occurrence dependence test proposed in Abbring et al. (2003), which we have applied on longitudinal data from France for the period 1995-1997.

Despite the short horizon of the panel, our results suggest the presence of moral hazard among inexperienced drivers with less than 15 years of experience. We do not find evidence of asymmetric learning for the vast majority of drivers. We find some evidence of asymmetric learning for those with less than 5 years of experience, which disappears quickly as both drivers and insurers learn about the underlying risks.

The results for the experienced group are largely consistent with the evidence presented in Abbring et al. (2003). For drivers with less experience, particularly those with less than 5 years of experience, we find evidence of both asymmetric learning and moral hazard, while they found only weak evidence of negative occurrence dependence. It is not clear what could explain the differences in our results for inexperienced drivers. On the one hand, we have used a longer horizon and more data concerning contracts, accidents

and the bonus-malus. On the other hand, our tests are highly parametric, which may lead to misspecification, whereas Abbring et al. (2003) constructed non-parametric tests, which may be less powerful when applied to rare events.

References

Abbring, J. H., P.A. Chiappori, and J. Pinquet (2003). "Moral Hazard and Dynamic Insurance Data," *Journal of the European Economic Association*, 1, pp. 767-820.

Abbring, J. H., P. A. Chiappori, and T. Zavadil (2008). "Better Safe than Sorry? Ex Ante and Ex Post Moral Hazard in Dynamic Insurance Data," mimeo, VU University of Amsterdam, 69 pages.

Alessie R., S. Hochguertel, and A. van Soest (2004). "Ownership of Stocks and Mutual Funds: a Panel Data Analysis," *The Review of Economics and Statistics*, 103, 6, pp. 783-796.

Attanasio, O. and G. Weber (1995). "Is Consumption Growth Consistent with Intertemporal Optimization? Evidence from the Consumer Expenditure Survey," *Journal of Political Economy*, 103, 6, pp. 1121-1157.

Boyer, M., G. Dionne and R. Kihlstrom (1989). "Insurance and the Value of Publicly Available Information," in *Studies in the Economics of Uncertainty: In Honor of Joseph Hadar*," edited by T. B. Fomby and T. K. Seo, New-York: Springer-Verlag, pp. 137-155.

Chiappori, P. A. (2000). "Econometric Models of Insurance Under Asymmetric Information," in *Handbook of Insurance*, edited by G. Dionne, Boston: Kluwer Academic Publishers, pp. 363-392.

Chiappori, P. A., F. Durand, and P. Y. Geoffard (1998). "Moral Hazard and the Demand for Physician Services: First Lessons from a French Natural Experiment," *European Economic Review*, 42, pp. 499-511.

Chiappori, P. A., B. Jullien, B. Salanie, and F. Salanie (2006). "Asymmetric Information in Insurance: General Testable Implications," *RAND Journal of Economics*, 37, 4, pp. 783-798.

Chiappori, P. A., I. Macho, P. Rey, and B. Salanie (1994). "Repeated Moral Hazard: The Role of Memory, Commitment, and the Access to Credit Markets," *European Economic Review*, 38, pp. 1527-1553.

Chiappori, P. A. and B. Salanie (2000). "Testing for Asymmetric Information in Insurance Markets," *Journal of Political Economy*, 108, 1, pp. 56-78.

Cohen, A. (2005). "Asymmetric Information and Learning: Evidence from the Automobile Insurance Market," *The Review of Economics and Statistics*, 87, 2, pp. 197-207.

Cohen, A. and P. Siegelman (2009). "Testing for Adverse Selection in Insurance Markets" *Journal of Risk and Insurance*, forthcoming.

Crocker, K. and A. Snow (1986), "The Efficiency Effects of Categorical Discrimination in the Insurance Industry," *Journal of Political Economy*, 94, pp. 321-344.

de Garidel-Thoron T. (2005). "Welfare-Improving Asymmetric Information in a Dynamic Insurance Markets," *Journal of Political Economy*, 113, 1, pp. 121-150.

Dionne, G. (2001). "Insurance Regulation in Other Industrial Countries," in *Deregulating Property-Liability Insurance*, edited by J. D. Cummins, Washington: AEI-Brookings Joint Center For Regulatory Studies, Washington, pp. 341-396.

Dionne, G. and N. Doherty (1994). "Adverse Selection, Commitment and Renegotiation: Extension to and Evidence from Insurance Markets," *Journal of Political Economy* 102, pp. 209-235.

Dionne, G., C. Gouieroux, and C. Vanasse (2001). "Testing for Evidence of Adverse Selection in the Automobile Insurance Market: A Comment," *Journal of Political Economy* 109, pp. 444-453.

Dionne, G., J. Pinquet, M. Maurice, and C. Vanasse (2005). "Point-Record Incentives, Asymmetric Information and Dynamic Data," mimeo, HEC Montreal, Canada Research Chair in Risk Management. 36 pages.

Dionne, G. and P. St-Michel (1991). "Workers' Compensation and Moral Hazard," *Review of Economics and Statistics*, 73, pp. 236-244.

Finkelstein, A. and K. McGarry (2006). "Multiple Dimensions of Private Information: Evidence from Long-term Care Insurance Market". *American Economic Review*, 96, 4, pp. 938-958.

Finkelstein, A. and J. Porteba (2004). Adverse Selection in Insurance Markets: Policyholders Evidence from the U. K. Annuity Market," *Journal of Political Economy*, 112, pp. 183-208.

Heckman J. J. (1981). "The Incidental Parameters Problem and the Problem of Initial Condition in Estimating a Discrete-Time Data Stochastic Process," in *Structural Analysis of Discrete Data with Econometric Applications*, edited by C.F. Manski and D. McFadden, Boston: MIT Press, pp. 179-195.

Hendel I. and A. Lizzeri (1999). "Adverse Selection in Durable Goods Markets," *American Economic Review*, 89, pp. 1097-1115.

Hendel I. and A. Lizzeri (2003). "The Role of Commitment in Dynamic Contracts: Evidence from Life Insurance," *Quarterly Journal of Economics*, 118, pp. 299-327.

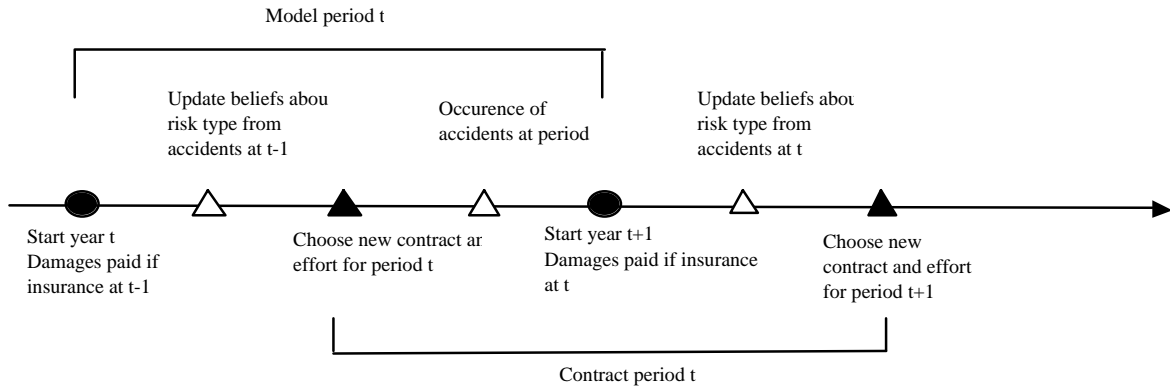
- Holmstrom, B. (1979). "Moral Hazard and Observability," *The Bell Journal of Economics*, 10, 1, pp. 74-92.
- Hurd, M.D. (1989). "Mortality Risk and Bequests," *Econometrica*, 57, 4, pp.779-813.
- Israel, M. (2004). "Do We Drive More Safely When Accidents are More Expensive? Identifying Moral Hazard from Experience Rating Schemes,' mimeo, Kellogg School of Management, Northwestern University, 23 p.
- Kim, H., D. Kim, and S. Im (2009). "Evidence of Asymmetric Information in the Automobile Insurance Market: Dichotomous Versus Multinomial Measurement of Insurance Coverage" *Journal of Risk and Insurance*, 76, 2, pp. 343-366.
- Lambert, R. A. (1983). "Long-Term Contracts and Moral Hazard" *Bell Journal of Economics*, 14, 2, pp. 441-452.
- Michaud P. -C. and K. Tatsiramos (2009). "Fertility and Female Employment Dynamics in Europe: The Effect of Using Alternative Econometric Modeling Assumptions," *Journal of Applied Econometrics*, forthcoming.
- Picard, P. (2000): "Les nouveaux enjeux de la régulation des marchés d'assurance," Working Paper no 2000-53, THEMA, Université Paris X-Nanterre. 27 pages
- Pinquet, J. (1999): "Une analyse des systèmes bonus-malus en assurance automobile", *Assurances* 67, 2, 241-249.
- Puelz R. and A. Snow (1994). "Evidence on Adverse Selection: Equilibrium Signaling and Cross Subsidization on the Insurance Market," *Journal of Political Economy* 102, pp. 236-257.
- Richaudeau R. (1999). "Automobile Insurance Contracts and Risk of Accident: an Empirical Test Using French Individual Data," *Geneva Papers on Risk and Insurance Theory*, 24, pp. 97-114.
- Rothschild M. and J. Stiglitz (1976). "Equilibrium in Insurance Markets: An Essay on the Economics of Imperfect Information," *Quarterly Journal of Economics*, 90, pp. 630-649.
- Schlesinger H. and J. M. G. von der Schulenburg (1993). "Consumer Information and Decisions to Switch Insurers," *The Journal of Risk and Insurance*, 60, 4, pp. 591-615
- Shavell S. (1979). "On Moral Hazard and Insurance," *Quarterly Journal of Economics*, November, 541-562.
- Wilson C. (1977). "A Model of Insurance Markets with Incomplete Information," *Journal of Economic Theory*, 16, pp. 167-207.

Winter R. (2000). "Moral Hazard in Insurance Markets," in *Handbook of Insurance*, Edited by G. Dionne, Boston: Kluwer Academic Publishers, pp.155-183.

Wooldridge, J.M. (2005). "A Simple Solution to the Initial Conditions Problem in Dynamic, Nonlinear Panel Data Models with Unobserved Heterogeneity," *Journal of Applied Econometrics*, 20, pp. 39-54.

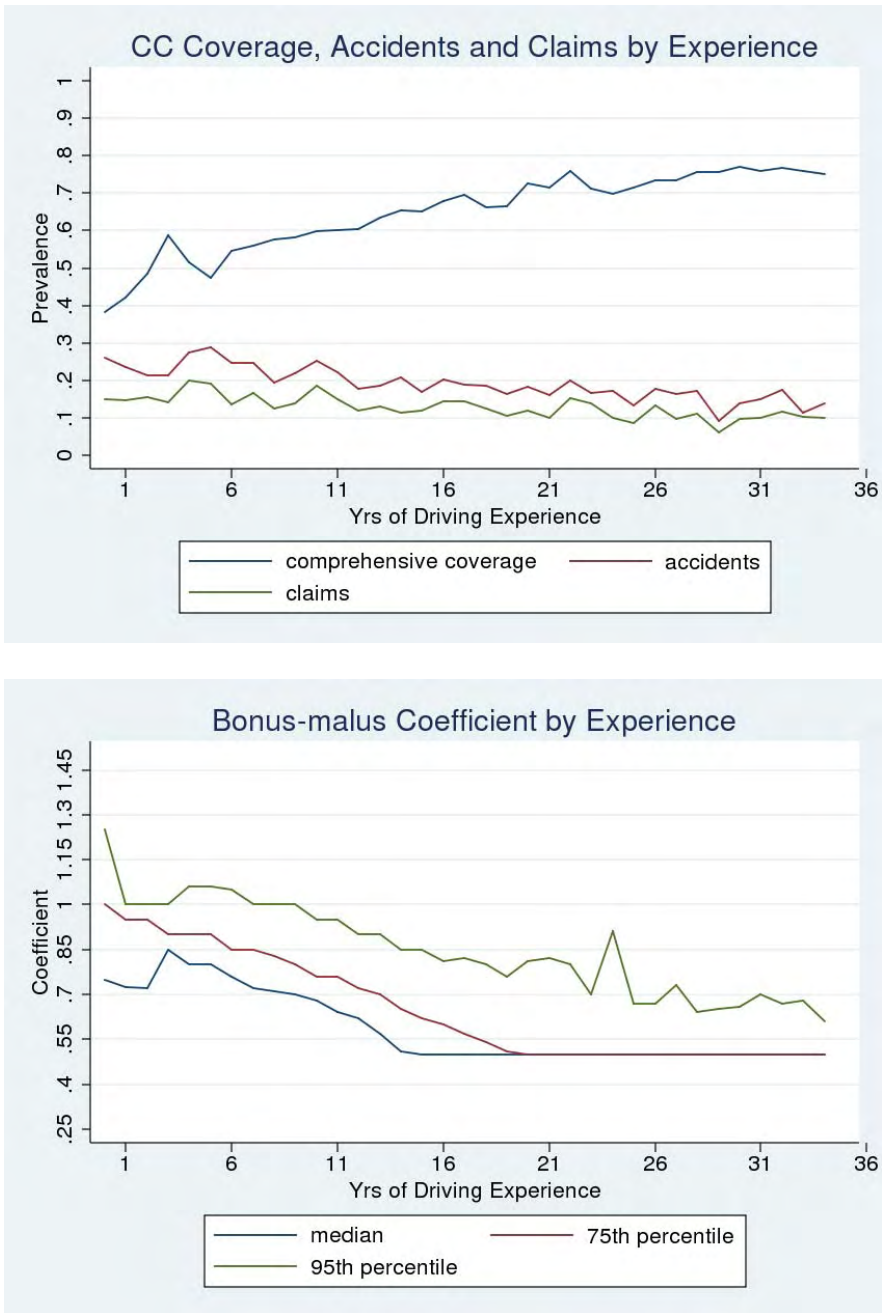
Tables and Figures

Figure 1 Timing of the Model and Contract



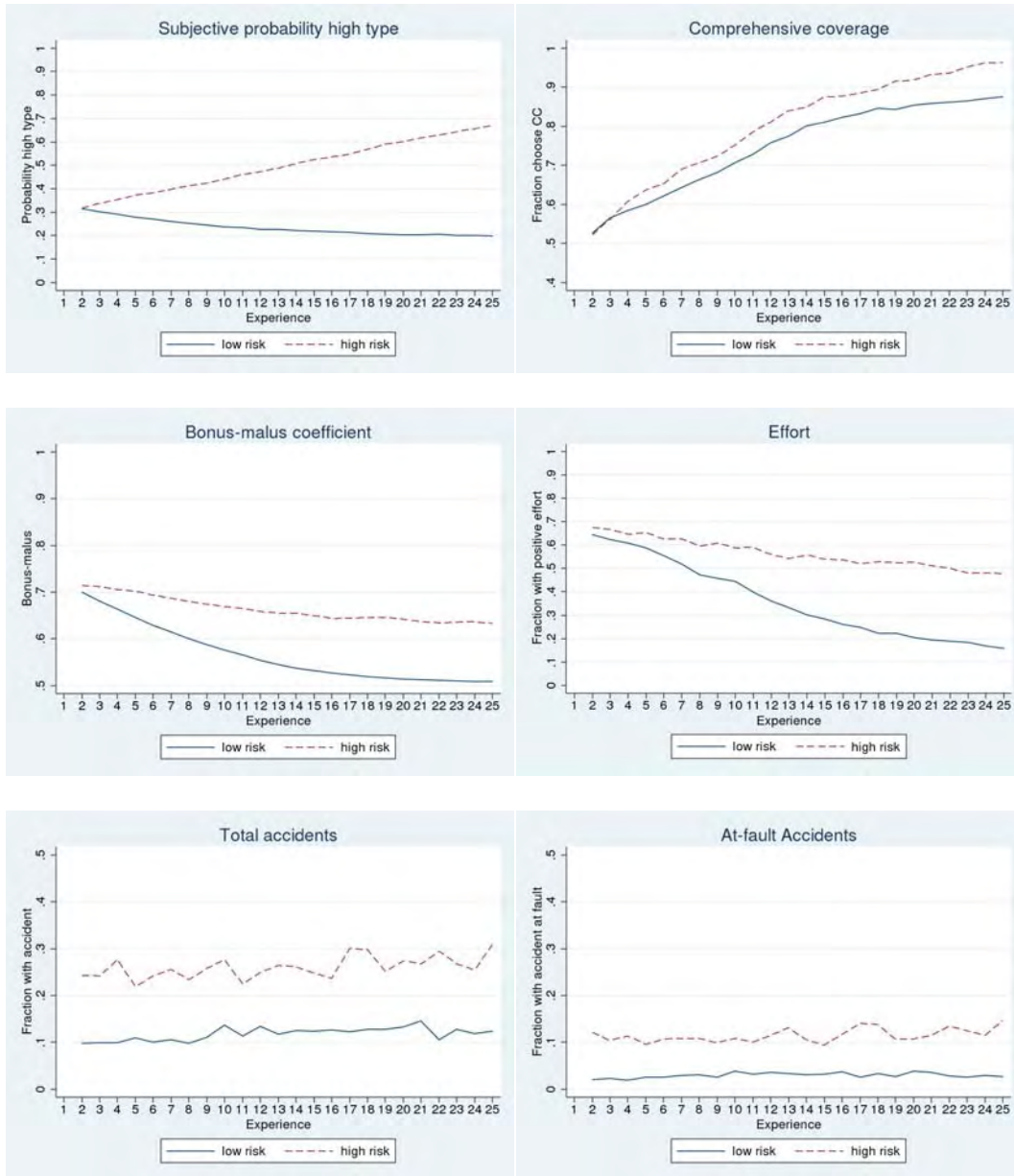
Notes: The figure shows the timing assumed in the model. Periods and contract do not coincide. In the model, individuals start by observing whether they had an accident or not; damages are paid by the insurer if an accident occurred. This is for a pre-determined contract which they have chosen in the previous year. Based on the occurrence of the accident but before renewing their contract, they update their beliefs on whether they are high risk or not. Based on the information they have at that point, they choose to renew or change their insurance contract. They also choose the effort level for the following year. This is the start of the contract period t . The model period t ends when an accident based on the effort level and insurance contract chosen in t occurs. In $t+1$, a new period model starts with the damages paid depending on the insurance contract and the occurrence of an accident. Beliefs are then revised and the individual renews the contract.

Figure 2 Profiles of Key Variables by Experience – SOFRES Panel



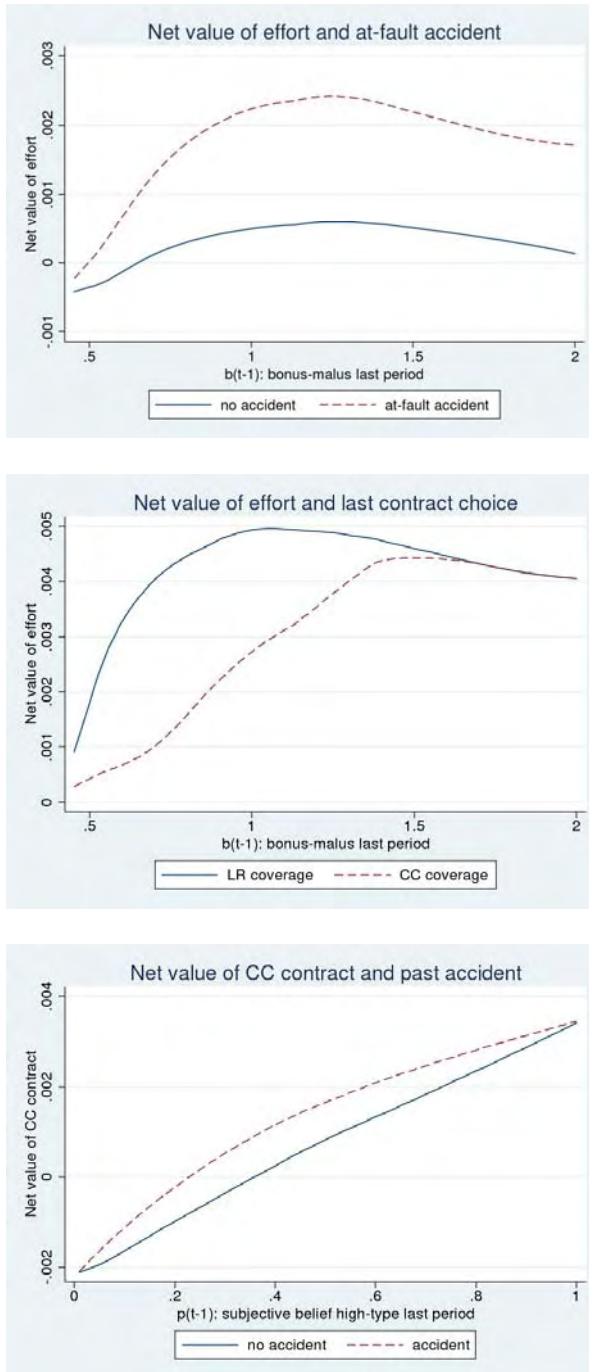
Source: Authors' calculations from the SOFRES panel. The top panel shows the prevalence of comprehensive coverage insurance contracts (CC), accidents (including claims) and claims by years of driving experience. The bottom panel reports the median, 75th percentile and 95th percentile of the bonus-malus coefficient distribution. The minimum value is 0.5 and more than 50% of drivers have a bonus of 0.5 after roughly 18 years.

Figure 3 Simulated Profiles of Key Variables under the Baseline Scenario with Moral Hazard and Asymmetric Learning



Notes: Statistics based on 2000 simulated individuals. Each profile is plotted by risk type.

Figure 4 Optimal Decision Rules, Moral Hazard and Asymmetric Learning



Notes: based on the numerical solution of the model. The net value of effort is defined as $\max(V^{11}, V^{01}) - \max(V^{10}, V^{00})$ and similarly for contract choice. Other variables in the state are set to $t=5$, $y=13,500$, $d(t-1)=0$, $e(t-1)=0$ and $p(t-1)=0.3$ for the top two graphs and $b(t-1)=0.75$ for the last one.

Table 1 Parameter Values

Parameter	Definition	Value	Source
σ	coefficient of relative risk aversion	1.5	literature
γ	marginal disutility effort	-5.00E-04	assumed
μ_{10}	constant, accident probability	-2.1	calibrated
μ_{11}	productivity effort, accident probability	-0.5	assumed
μ_{12}	factor high risk type, accident probability	0.75	assumed
μ_{20}	constant, at fault probability	-3.2	calibrated
μ_{21}	productivity effort, at fault probability	-0.5	assumed
μ_{22}	factor high risk, at fault probability	1.6	assumed
f	deductible	0.192	FFSA
L	loss	2.439	FFSA
ρ_0	log premium LR	-1.69	FFSA
ρ_1	log % change in premium for CC contract	0.7	FFSA
β	discount factor	0.985	assumed
δ_b	bonus factor	0.95	FFSA
δ_m	malus factor	1.25	FFSA
ψ	switching or search cost (euros)	0.05	assumed
δ	fraction of population high risk	0.3	assumed

Notes: parameters of the models along with the value used in the baseline simulations.

The source for the value is also reported. Calibrated means that the parameter was chosen to match certain features of the data as defined in text. Assumed implies that values were chosen arbitrarily. FFSA stands for the "Federation Francaise des Societes d'Assurances".

Table 2 Tests of Asymmetric Information: Moral Hazard and Asymmetric Learning

MH1	MH2	AL	
		no effect	effect
no effect	no effect	No MH or AL	AL
	effect	MH	AL and MH
effect	MH2		
	no effect	MH	AL
	effect	MH	AL and MH

Notes: The MH1 test involves the effect of the bonus-malus on accidents. The MH2 tests involves the effect of the previous contract choice on accidents. Finally, the AL test looks at the effect of accidents on contract choice. The table shows the conclusions that can be reached based on the results of the tests. Abbring et al. (2003) find an effect under MH1 for young drivers but cannot attribute it to asymmetric learning or moral hazard (no MH or AL). The three tests together allow us to separate the two information problems.

Table 3 Empirical Tests on Simulated Data: Information Problems

	Underlying DGP			
	MH		No MH	
	AL	No AL	AL	No AL
Test on Contract Choice				
n(t-1) - AL	0.428 (0.001)	-0.062 (0.327)	0.684 (0.001)	-0.261 (0.001)
Test on Accidents				
b(t) - MH1	-0.575 (0.001)	-0.674 (0.001)	-0.063 (0.537)	-0.133 (0.131)
d(t-1) - MH2	0.116 (0.003)	0.1344 (0.001)	-0.131 (0.013)	-0.0529 -0.407

Notes: bivariate probit coefficient estimates along with p-values. Estimates obtained by maximum simulated likelihood with 50 draws per respondent and equation. Different specifications are based on simulated data from varying data generating processes (DGP). The first column assumes both moral hazard (MH) and asymmetric learning (AL). The other three specifications vary the presence of information problems. Each specification controls for experience and income and deals with the initial condition problem as mentioned in the text. Complete results available upon request.

Table 4 Empirical Tests on Simulated Data: Initial Conditions and Unobserved Heterogeneity

	Specification			
	Wooldridge	Control α	Control α and $\pi(t-1)$	No control
Test on Contract Choice				
n(t-1) - AL	0.428 (0.001)	0.421 (0.001)	0.894 (0.001)	0.456 (0.001)
Test on Accidents				
b(t) - MH1	-0.575 (0.001)	-0.424 (0.001)	-0.294 (0.001)	0.109 (0.159)
d(t-1) - MH2	0.116 (0.003)	0.095 (0.001)	0.1577 (0.001)	0.176 (0.001)

Notes: bivariate probit coefficient estimates along with p-values. Each column consists of a different specification. The first column implements the Wooldridge solution to the initial condition problem and integrates out remaining unobserved heterogeneity by simulation (50 draws). The second specification controls directly for the risk type. The third specification adds the lag subjective probability of being high risk type. The specification in the last column does not control for unobserved heterogeneity nor corrects for left-censoring. Each specification controls for experience and income. Complete results available upon request.

Table 5 Tests on SOFRES Panel

Test on Contract Choice	Years of Experience Group		
	Less than 15 years	Interaction < 5 yrs	More than 15 years
n(t-1) - AL Test	0.152 (0.729)	0.034 (0.940)	-0.223 (0.483)
n(t-1) x exp<5		0.868 (0.053)	
Test on Accidents			
b(t-1) - MH1 Test	-2.241 (0.081)	-2.216 (0.080)	0.096 (0.892)
d(t-1) - MH2 Test	0.651 (0.026)	0.622 (0.033)	-0.14 (0.661)
d(t-1) x exp<5		0.170 (0.351)	
Sample Size (N x T)	1066	1066	1537

Notes: bivariate probit coefficient estimates along with p-values based on robust standard errors. Estimates obtained from maximum simulated likelihood with 100 draws per respondent and equation. The first column reports coefficients estimated on drivers with less than 15 years of experience. The second column allows for different AL and MH2 tests for those with less than 5 years of experience in the contract choice equation. The last column presents the coefficients estimated only on those with more than 15 years of experience. Other controls include experience, income and age of vehicle as well as initial conditions. Complete results available upon request.

Appendix

Data

The survey conducted by SOFRES is called *Parc Automobile*. It is part of a larger monthly consumer survey. Monthly surveys are targeted to a sub-sample of a stable pool of approximately 20,000 respondents. Each year since 1983, the January survey targets roughly 10,000 respondents to answer questions regarding the use of their vehicles. The sample used for the *Parc Automobile* survey is renewed each year so that it is representative of the French population. Each year, 1/3 of the respondents are not re-interviewed, making the *Parc Automobile* survey a rotating panel. Until 1995, the survey did not contain much information on accidents and insurance coverage, so we use data from 1995 to 1997 in our analysis. The respondent is the head of the household and questions pertaining to socio-demographic characteristics (i.e. gender) pertain only to the respondent.

The dataset provided by SOFRES consists of three files containing data for each year (1995, 1996 and 1997) on policyholders and their vehicles. Each observation represents a vehicle-policyholder pair. We delete observations where the vehicle has not been used (zero kilometers reported) or where information is missing on key variables (e.g. accidents, bonus-malus, comprehensive coverage). For example, the potential sample from 1997 consists of 8849 observations. Eliminating those with zero kilometers yields 7659 observations. Finally, dropping observations where key variables are missing leaves us with 5279 observations for that year.

Although the dataset contains an identifier for each respondent that is consistent across years, it does not contain a unique identifier for vehicles. However, we have access to the first 4 digits of the license plate number and the year the vehicle was built, which we use to merge the datasets for each year in combination with the respondents' ID. In this process, we drop observations where the license plate and/or year of manufacturing are missing or incomplete. For 1997, we are left with 5121 observations. To validate our match, we have checked whether the characteristics of vehicles that should not change across years (such as make, vehicle class and fuel type) were indeed constant. When they changed, we classified the observation as an invalid match and dropped the entire record. In the end, we were left with 4960 observations for analysis. The same pattern of exclusion was applied to other years with a similar proportion of observations dropped. Table A.1 reports the record number per year stratified by the number of years in the panel. Overall, 1049 observations are present in all three years. Table A.2 reports descriptive statistics on some of the variables used in the analysis, including socio-demographic characteristics of respondents.

Table A.1 Structure of SOFRES Panel

Years in Panel	Year of Survey		
	1995	1996	1997
1 year	3052	2454	2901
2 years	1342	2362	1010
3 years	1049	1049	1049
Total	5443	5855	4960

Notes: Each cell gives the number of records stratified by how many years the policyholder-vehicle pair is in the panel.

Table A.2 Descriptive Statistics on Panel

	%		%
Gender (male)	62.9	Region	
Age		Paris	33.4
18-24	5.2	North	9.2
25-34	22.3	East	9.8
35-44	21.7	South	35.2
45-54	16.5	West	12.4
55-64	14.7	Primary Use	
65+	20	Rural	13.3
Occupation		City	40.6
Retired	28.3	Mixed	46.6
Farmer	4.3	Age of vehicle	
Manager	8.0	< 3 years	34.3
Professional	7.6	3-5 yrs	18.6
Teacher (active or retired)	19.4	5-10 yrs	33.4
Employee	29.1	10+ yrs	13.7
Student	3.2	Experience	
Number of vehicles		25th pct.	14
1	45.8	Median	24
2	45.6	75th pct.	36
3+	8.5	Mean	25.2

Notes: statistics for the 1995 wave of the survey.

Table A.3 Transitions by Experience Level

fraction	Years of Experience	
	<15 years	> 15 years
Accidents		
1 -> 1	0.295	0.192
1 -> 0	0.705	0.802
0 -> 1	0.147	0.123
0 -> 0	0.852	0.877
% accident	0.216	0.167
CC coverage		
1 -> 1	0.916	0.944
1 -> 0	0.084	0.057
0 -> 1	0.030	0.066
0 -> 0	0.969	0.934
% CC	0.586	0.718

Notes: Authors' calculations using all years (1995-1997). Transition rates from origin x at t-1 to y at t are reported as x -> y where 0 denotes no accident or CC coverage and 1 an accident or CC coverage.